RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _____ 09/526,298 _____

Source: _____ JFW16, _____

Date Processed by STIC: _____ 05/30/2006 _____

ENTERED

> **This Raw Listing contains the General Information Section and up to the first 5 pages.**

```
 1                                  SEQUENCE LISTING
 2
 3    (1)      General Information:
 4
 5        (i) APPLICANT: EVANS, RONALD M.
 6                       MCKEOWN, MICHAEL B.
 7                       ORO, ANTHONY E.
 8                       SEGRAVES, WILLIAM A.
 9                       YAO, TSO-PANG
10
11       (ii) TITLE OF INVENTION: MULTIMERIC FORMS OF MEMBERS OF THE
12               STEROID/THYROID SUPERFAMILY OF RECEPTORS WITH THE
13               ULTRASPIRACLE RECEPTOR
14
15      (iii) NUMBER OF SEQUENCES: 29
16
17       (iv) CORRESPONDENCE ADDRESS:
18               (A) ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
19               (B) STREET: 444 South Flower Street, Suite 2000
20               (C) CITY: Los Angeles
21               (D) STATE: California
22               (E) COUNTRY: United States
23               (F) ZIP: 90071
24
25        (v) COMPUTER READABLE FORM:
26               (A) MEDIUM TYPE: Floppy disk
27               (B) COMPUTER: IBM PC compatible
28               (C) OPERATING SYSTEM: PC-DOS/MS-DOS
29               (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
30
31       (vi) CURRENT APPLICATION DATA:
32               (A) APPLICATION NUMBER: US 09/526,298
33               (B) FILING DATE: 15-MARCH-2000
34               (C) CLASSIFICATION:
35      (vii) PRIOR APPLICATION DATA:
36               (A) APPLICATION NUMBER: US 07/907,908
37               (B) FILING DATE: 02-JUL-1992
38               (C) CLASSIFICATION:
39     (viii) ATTORNEY/AGENT INFORMATION:
40               (A) NAME: Reiter,, Stephen E.
41               (B) REGISTRATION NUMBER: 31192
42               (C) REFERENCE/DOCKET NUMBER: P41 9321
43
44       (ix) TELECOMMUNICATION INFORMATION:
45               (A) TELEPHONE: (619) 546-4737
46               (B) TELEFAX: (619) 546-9392
```

INPUT SET: S30829.raw

```
47
48
49     (2) INFORMATION FOR SEQ ID NO:1:
50
51         (i) SEQUENCE CHARACTERISTICS:
52                (A) LENGTH: 2304 base pairs
53                (B) TYPE: nucleic acid
54                (C) STRANDEDNESS: single
55                (D) TOPOLOGY: linear
56
57         (ii) MOLECULE TYPE: DNA (genomic)
58
59
60         (ix) FEATURE:
61                (A) NAME/KEY: CDS
62                (B) LOCATION: 163..1701
63
64
65         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
66
67     GGACACGGTG GCGTTGGCAA AGTGAAACCC CAACAGAGAG GCGAAAGCGA GCCAAGACAC       60
68
69     ACCACATACA CACGAAGAGA ACGAGCAAGA AGAAACCGGT AGGCGGAGGA GGCGCTGCCC      120
70
71     CCAGTTCCTC CAATATACCC AGCACCACAT CACAAGCCCA GG ATG GAC AAC TGC        174
72                                                  Met Asp Asn Cys
73                                                   1
74
75     GAC CAG GAC GCC AGC TTT CGG CTG AGC CAC ATC AAG GAG GAG GTC AAG      222
76     Asp Gln Asp Ala Ser Phe Arg Leu Ser His Ile Lys Glu Glu Val Lys
77      5                   10                  15                  20
78
79     CCG GAC ATC TCG CAG CTG AAC GAC AGC AAC AAC AGC AGC TTT TCG CCC      270
80     Pro Asp Ile Ser Gln Leu Asn Asp Ser Asn Asn Ser Ser Phe Ser Pro
81                      25                  30                  35
82
83     AAG GCC GAG AGT CCC GTG CCC TTC ATG CAG GCC ATG TCC ATG GTC CAC      318
84     Lys Ala Glu Ser Pro Val Pro Phe Met Gln Ala Met Ser Met Val His
85                  40                  45                  50
86
87     GTG CTG CCC GGC TCC AAC TCC GCC AGC TCC AAC AAC AAC AGC GCT GGA      366
88     Val Leu Pro Gly Ser Asn Ser Ala Ser Ser Asn Asn Asn Ser Ala Gly
89              55                  60                  65
90
91     GAT GCC CAA ATG GCG CAG GCG CCC AAT TCG GCT GGA GGC TCT GCC GCC      414
92     Asp Ala Gln Met Ala Gln Ala Pro Asn Ser Ala Gly Gly Ser Ala Ala
93          70                  75                  80
94
95     GCT GCA GTC CAG CAG CAG TAT CCG CCT AAC CAT CCG CTG AGC GGC AGC      462
96     Ala Ala Val Gln Gln Gln Tyr Pro Pro Asn His Pro Leu Ser Gly Ser
97      85                  90                  95                  100
98
99     AAG CAC CTC TGC TCT ATT TGC GGG GAT CGG GCC AGT GGC AAG CAC TAC      510
```

INPUT SET: S30829.raw

```
100   Lys His Leu Cys Ser Ile Cys Gly Asp Arg Ala Ser Gly Lys His Tyr
101                   105             110             115
102
103   GGC GTG TAC AGC TGT GAG GGC TGC AAG GGC TTC TTT AAA CGC ACA GTG      558
104   Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val
105                   120             125             130
106
107   CGC AAG GAT CTC ACA TAC GCT TGC AGG GAG AAC CGC AAC TGC ATC ATA      606
108   Arg Lys Asp Leu Thr Tyr Ala Cys Arg Glu Asn Arg Asn Cys Ile Ile
109           135             140             145
110
111   GAC AAG CGG CAG AGG AAC CGC TGC CAG TAC TGC CGC TAC CAG AAG TGC      654
112   Asp Lys Arg Gln Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys
113           150             155             160
114
115   CTA ACC TGC GGC ATG AAG CGC GAA GCG GTC CAG GAG GAG CGT CAA CGC      702
116   Leu Thr Cys Gly Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg
117   165             170             175             180
118
119   GGC GCC CGC AAT GCG GCG GGT AGG CTC AGC GCC AGC GGA GGC GGC AGT      750
120   Gly Ala Arg Asn Ala Ala Gly Arg Leu Ser Ala Ser Gly Gly Gly Ser
121                   185             190             195
122
123   AGC GGT CCA GGT TCG GTA GGC GGA TCC AGC TCT CAA GGC GGA GGA GGA      798
124   Ser Gly Pro Gly Ser Val Gly Gly Ser Ser Ser Gln Gly Gly Gly Gly
125               200             205             210
126
127   GGA GGC GGC GTT TCT GGC GGA ATG GGC AGC GGC AAC GGT TCT GAT GAC      846
128   Gly Gly Gly Val Ser Gly Gly Met Gly Ser Gly Asn Gly Ser Asp Asp
129           215             220             225
130
131   TTC ATG ACC AAT AGC GTG TCC AGG GAT TTC TCG ATC GAG CGC ATC ATA      894
132   Phe Met Thr Asn Ser Val Ser Arg Asp Phe Ser Ile Glu Arg Ile Ile
133           230             235             240
134
135   GAG GCC GAG CAG CGA GCG GAG ACC CAA TGC GGC GAT CGT GCA CTG ACG      942
136   Glu Ala Glu Gln Arg Ala Glu Thr Gln Cys Gly Asp Arg Ala Leu Thr
137   245             250             255             260
138
139   TTC CTG CGC GTT GGT CCC TAT TCC ACA GTC CAG CCG GAC TAC AAG GGT      990
140   Phe Leu Arg Val Gly Pro Tyr Ser Thr Val Gln Pro Asp Tyr Lys Gly
141                   265             270             275
142
143   GCC GTG TCG GCC CTG TGC CAA GTG GTC AAC AAA CAG CTC TTC CAG ATG     1038
144   Ala Val Ser Ala Leu Cys Gln Val Val Asn Lys Gln Leu Phe Gln Met
145               280             285             290
146
147   GTC GAA TAC GCG CGC ATG ATG CCG CAC TTT GCC CAG GTG CCG CTG GAC     1086
148   Val Glu Tyr Ala Arg Met Met Pro His Phe Ala Gln Val Pro Leu Asp
149           295             300             305
150
151   GAC CAG GTG ATT CTG CTG AAA GCC GCT TGG ATC GAG CTG CTC ATT GCG     1134
152   Asp Gln Val Ile Leu Leu Lys Ala Ala Trp Ile Glu Leu Leu Ile Ala
```

INPUT SET: S30829.raw

```
153        310                   315                   320
154
155  AAC GTG GCC TGG TGC AGC ATC GTT TCG CTG GAT GAC GGC GGT GCC GGC    1182
156  Asn Val Ala Trp Cys Ser Ile Val Ser Leu Asp Asp Gly Gly Ala Gly
157  325                   330                   335                   340
158
159  GGC GGG GGC GGT GGA CTA GGC CAC GAT GGC TCC TTT GAG CGA CGA TCA    1230
160  Gly Gly Gly Gly Gly Leu Gly His Asp Gly Ser Phe Glu Arg Arg Ser
161                        345                   350                   355
162
163  CCG GGC CTT CAG CCC CAG CAG CTG TTC CTC AAC CAG AGC TTC TCG TAC    1278
164  Pro Gly Leu Gln Pro Gln Gln Leu Phe Leu Asn Gln Ser Phe Ser Tyr
165                        360                   365                   370
166
167  CAT CGC AAC AGT GCG ATC AAA GCC GGT GTG TCA GCC ATC TTC GAC CGC    1326
168  His Arg Asn Ser Ala Ile Lys Ala Gly Val Ser Ala Ile Phe Asp Arg
169                   375                   380                   385
170
171  ATA TTG TCG GAG CTG AGT GTA AAG ATG AAG CGG CTG AAT CTC GAC CGA    1374
172  Ile Leu Ser Glu Leu Ser Val Lys Met Lys Arg Leu Asn Leu Asp Arg
173       390                   395                   400
174
175  CGC GAG CTG TCC TGC TTG AAG GCC ATC ATA CTG TAC AAC CCG GAC ATA    1422
176  Arg Glu Leu Ser Cys Leu Lys Ala Ile Ile Leu Tyr Asn Pro Asp Ile
177  405                   410                   415                   420
178
179  CGC GGG ATC AAG AGC CGG GCG GAG ATC GAG ATG TGC CGC GAG AAG GTG    1470
180  Arg Gly Ile Lys Ser Arg Ala Glu Ile Glu Met Cys Arg Glu Lys Val
181                        425                   430                   435
182
183  TAC GCT TGC CTG GAC GAG CAC TGC CGC CTG GAA CAT CCG GGC GAC GAT    1518
184  Tyr Ala Cys Leu Asp Glu His Cys Arg Leu Glu His Pro Gly Asp Asp
185                   440                   445                   450
186
187  GGA CGC TTT GCG CAA CTG CTG CTG CGT CTG CGC CGC TTT GCG ATC GAT    1566
188  Gly Arg Phe Ala Gln Leu Leu Leu Arg Leu Arg Arg Phe Ala Ile Asp
189       455                   460                   465
190
191  CAG CCT GAA GTG CCA GGA TCA CCT GTT CCT CTT CCG CAT TAC CAG CGA    1614
192  Gln Pro Glu Val Pro Gly Ser Pro Val Pro Leu Pro His Tyr Gln Arg
193       470                   475                   480
194
195  CCG GCC GCT GGA GGA GCT CTT TCT CGA GCA GCT GGA GGC GCC GCC GCC    1662
196  Pro Ala Ala Gly Gly Ala Leu Ser Arg Ala Ala Gly Gly Ala Ala Ala
197  485                   490                   495                   500
198
199  ACC CGG CCT GGC GAT GAA ACT GGA GTA GGG TCC CGA CTC TAAAGTCGCC    1711
200  Thr Arg Pro Gly Asp Glu Thr Gly Val Gly Ser Arg Leu
201                   505                   510
202
203  CCCGTTCTCC ATCCGAAAAA TGTTTCATTG TGATTGCGTT TGTTTGCATT TCTCCTCTCT   1771
204
205  ATCCCTACAA AAGCCCCCTA ATATTACGCA AAATGTGTAT GTAATTGTTT ATTTTTTTT    1831
```

206
207   TATTACCTAA TATTATTATT ATTATTGATA TAGAAAATGT TTTCCTTAAG ATGAAGATTA   1891
208
209   GCCTCCTCGA CGTTTATGTC CCAGTAAACG AAAAACAAAC AAAATCCAAA ACTTGAAAAG   1951
210
211   AACACAAAAC ACGAACGAGA AAATGCACAC AAGCAAAGTA AAAGTAAAAG TTAAACTAAA   2011
212
213   GCTAAACGAG TAAAGATATT AAAATAACGG TTAAAATTAA TGCATAGTTA TGATCTACAG   2071
214
215   ACGTATGTAA ACATACAAAT TCAGCATAAA TATATATGTC AGCAGGCGCA TATCTGCGGT   2131
216
217   GCTGGCCCCG TTCTAAACCA ATTGTAATTA CTTTTTAACA TAAATTTACC CAAAACGTTA   2191
218
219   TCAATTAGAT GCGAGATACA AAAATCACCG ACGAAACCA ACAAAATATA TCTATGTATA   2251
220
221   AAAAATATAA GCTGCATAAC AAAAAAAAAA AAAAAAAAAA AAAAAAAAA AAA   2304
222
223
224   (2) INFORMATION FOR SEQ ID NO:2:
225
226   (i) SEQUENCE CHARACTERISTICS:
227   (A) LENGTH: 513 amino acids
228   (B) TYPE: amino acid
229   (D) TOPOLOGY: linear
230
231   (ii) MOLECULE TYPE: protein
232
233   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
234
235   Met Asp Asn Cys Asp Gln Asp Ala Ser Phe Arg Leu Ser His Ile Lys
236   1   5   10   15
237
238   Glu Glu Val Lys Pro Asp Ile Ser Gln Leu Asn Asp Ser Asn Asn Ser
239   20   25   30
240
241   Ser Phe Ser Pro Lys Ala Glu Ser Pro Val Pro Phe Met Gln Ala Met
242   35   40   45
243
244   Ser Met Val His Val Leu Pro Gly Ser Asn Ser Ala Ser Ser Asn Asn
245   50   55   60
246
247   Asn Ser Ala Gly Asp Ala Gln Met Ala Gln Ala Pro Asn Ser Ala Gly
248   65   70   75   80
249
250   Gly Ser Ala Ala Ala Ala Val Gln Gln Gln Tyr Pro Pro Asn His Pro
251   85   90   95
252
253   Leu Ser Gly Ser Lys His Leu Cys Ser Ile Cys Gly Asp Arg Ala Ser
254   100   105   110
255
256   Gly Lys His Tyr Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly Phe Phe
257   115   120   125
258